2 October 1963

RSIC-92

AUTOMATION PROGRESS AT RSIC

THE STATUS OF ALPHA 1

By

F. E. Croxton

Redstone Scientific Information Center
U S ARMY MISSILE COMMAND
REDSTONE ARSENAL, ALABAMA

OK CFSTI

AUTOMATION PROGRESS AT RSIC

THE STATUS OF ALPHA 1

Prepared for Delivery Before the 7th Annual Military Librarian's Workshop
2 October 1963 White Oak, Maryland

When I attended my first military librarian's workshop last year at White Sands, I did not expect to be given the opportunity to speak to you so soon. However, Mr. Vlannes asked that I cover for you some of the automation work we are doing at Redstone Arsenal.

The few minutes that I will take will be concerned with a progress report on the system to which we have given the acronym, ALPHA I. In our development of ALPHA I, we have attempted to view our library-type activities and our information retrieval efforts as an organic whole and to devise an integrated method which will allow us to solve at one time both the more or less administrative aspects of running a library, as well as some of the more sophisticated subjective aspects of answering questions. We have dubbed our current condition a first generation study because we know that even before it is completed, we will be trying to improve the smoothness and sophistication of both our methods and our machine techniques. In developing ALPHA 1, which differs from any other method with which I am familiar only in that it is a truly integrated study, we have undertaken the work in the same way that any good development program would be undertaken. We assembled an interdisciplinary team of people who could bring to bear on this problem all of the necessary skills -- in this case, scientific skills, library skills, and computer system analysis skills.

SLIDE

Like many government projects, this is being done with both in-house and contractor participation. Specifically, it is a joint activity of the Redstone Scientific Information Center, the Computation Center of the Army Missile Support Command, and the Computer Division of the General Electric Company, contractors for the operation of the Computer Laboratory of Marshall Space Flight Center, NASA. Members of each of the three groups have made both general and specific contributions.

Let me get the question of contributions and credits settled to begin with - we have borrowed information from everyone that we found who had what we felt was a good idea. We happily credit Picatinny Arsenal, IBM Corporation, Eli Lilly, Sandia Corporation, University of California, and General Electric's previous work at Western Reserve, University of Chicago, and MEDLARS. If we have used ideas originated by anyone else, I hope they will let me know so I can include their names in any future presentation.

Before discussing ALPHA I, a few words about our activity at Redstone Arsenal, will help you to visualize the sort of problem we face and to explain briefly why we were not only interested in automation, but were actually forced to it. Redstone Scientific Information Center serves both the Army Missile Command and the Marshall Space Flight Center. The potential patron group approaches 20,000 employees. In addition, the on-site and local contractor organizations, and others who call upon us for service, add nearly another 50% to our patron group. We estimate that we serve 6,000 scientific and engineer professionals spread over an area of slightly more than 100 square miles.

Their requirements for information service increased by over 50% from last summer to this spring. Naturally, manpower cannot be added at this rate. In fact, it has been necessary, due in part to the Army reorganization which took place last year and to other factors which decreased the number of employees available in our command, to decrease the number of people performing library functions in spite of the heavy increase in service requirements - quality naturally may have suffered.

Our long-term answer to this problem is ALPHA I. As we approached the study of ALPHA, our first problem was to define a library. This was not too difficult for we feel a library can be considered to be a communication link.

SLIDE A link through which a worker of the past speaks to one in the present, or a worker of the present conveys his thoughts to anyone who may require them. As we look at this diagram, some of the gross characteristics which ALPHA must have become evident. We really require only information about the patron, and about the holdings, whether they be books, periodicals, documents, or all three.

After our systems analysis of what was actually done in the library was completed, it was apparent that the traditional distinction in libraries between these three types of material continued to be appropriate in mechanized systems. There were significant differences, particularly between periodicals and the other two types. We are, therefore, devising a system

SLIDE which is composed of 4 basic parts: A master file of information about our patron and master files of information about the three primary kinds of items

available to him. Of course, we have more than books, documents, and periodicals, and although the thoughts and illustrations here can be replicated for each kind of material, the other kinds normally can be forced into the pattern for one of the three represented on the slide.

You will notice we have placed the patron file in the center of this very simple diagram. I will not belabor you with the obvious reasons for the patron being in the central position in all of our efforts because they are self-evident to all of you service-oriented librarians.

You will notice that there are two levels of tapes surrounding our patron and by the way, these shapes represent magnetic tapes in the symbology of the systems analyst and the computer man. The inner group of tapes are those prepared from the masters to do jobs more efficiently, or prepared prior to entry of the data on the masters. We will look at each of them in a moment, but first let us examine what data appear in the master files.

SLIDE

This is a generalized representation of the patron tape record showing the identifying and descriptive data and the requirements and interest profiles of each of our customers. This file will be arranged in the order of social security number. You may recall that I mentioned that we serve several organizations at Redstone. Payroll numbers, badge numbers, and similar local identifications do not furnish us with unique identifiers; hence the need for the social security number. It is accidental in a sense that we have selected one which can also be used in any other locality.

SLIDE Here we see the gross arrangement of data on the periodical master record. We have information regarding the magazine itself, the publishing and cost aspects required for reordering, the predetermined binding information, the holdings record, and the essential reciprocal of that, the list of lacks.

SLIDE Here is an indication of the general format of the book master; again information about the book itself, the call number and bibliographic string, the housekeeping information, i.e., copies available and descriptive information, then the more complex data, such as the subject entries, etc.

SLIDE The document master is essentially in the same format as the book master; however, since there are some specific differences in the elements of information and in some parts of the processing, it appears appropriate to keep documents separated from books at this time.

In developing ALPHA I, seven almost axiomatic principles were borne in

SLIDE mind. These were:

1. Machine readability shall be obtained at the earliest possible time. In general, this is done at the time of ordering or receiving.

2. Redundant transcription shall be minimized by using prepunched transaction cards, i.e., feedback shall be used to the maximum extent possible.

3. Generalizability is required to provide for possible changes in our requirements. We may grow or we may shrink. We may change emphasis -- our system must change with us.

4. Open-endedness and provisions for modification are essential if we are to take advantage of the efforts of others like NASA, DDC, AEC, or any other groups - local or central - who automate their information collections. What I mean to say is that we must have active nerve endings.

5. Many types of transactions are required to maintain the master and subsidiary files. The simplest external technique is the hopper method in which all transaction cards are thrown together and the machine makes all processing decisions.

6 We will use the work of others wherever we can to avoid unnecessary expense and to speed our own efforts.

7. Each automated process must result in at least as satisfactory a manual tool as was available through manual methods in addition to a machine interrogatable file.

You may be able to see, as I show you a few more diagrams, how some of these principles - particularly the one of feedback - are used wherever possible.

SLIDE Let us take the book-related activities first. Now we have identified the inner group of tapes. The in-process tape is used in conjunction with each of the three master tapes - books, documents, and periodicals. The inventory tape is used for both books and documents. Let us follow this diagram for a moment or two to see how things will work Let us assume that a patron needs a book. His request is first processed against the inventory file (in practice to the shelf) to determine whether a copy is available. This file, in many ways, is a combination of the traditional shelf list and the traditional circulation record -- it has an entry for every copy of every title.

If a copy is available, it is loaned to him and the loan information is posted to this file. A suitable extraction of this file is a circulation report arranged both by call number and by borrower's name. Overdue actions can be triggered.

If the book is not available for circulation, the in-process file is interrogated to determine whether there is an uncommitted copy on order, if not, recall or reorder can be initiated. In recall, the inventory file is posted. If it is the librarian's decision to order, the order data are posted to the in-process file. Periodically, the in-process file is summarized from a financial standpoint.

The on-order list, which is a listing of the in-process file, has flags showing those actually received but not yet entered in the bibliography or inventory file. Upon receipt of a book, the flag is added to the in-process file showing that the item is awaiting cataloging. When this is completed, information is transferred, in the case of an added copy, to the inventory master or of a new title, to the bibliography master as well. Then, as part of this file maintenance process on the bibliography master, an accession list will be produced and selective dissemination accomplished on the basis of data on the patron file. Multiple sequencing of the bibliography master, of course, will result in printed lists corresponding to the traditional catalog which, like any book catalog, can be available for use in many locations. Information retrieval searches will be performed against an inverted entry tape produced from this bibliography file, using a specially developed infinite query search algorithm. In this way, with a minimum number

of tapes, all of the material which we feel is required to run an effective operation is produced. This includes data which indicates the position of an item on loan and which has overdue recall built into it, catalogs to show what material is available, management information, and a positiv current dissemination technique.

SLIDE Now, let us examine the document related activities. You will notice that in many ways this appears to be a mirror image of what you just saw. This is as it should be for the same operations are involved for both types of material; the same functions must be performed. There are certain differences, however, in the techniques we will use. For example, our book circulation method depends partially on the flow of a card with the book which is used to discharge the book upon return, while our corresponding document circulation method will retain this card in a manual file in the library, signed as a receipt when necessary, prior to its use as a discharging mechanism. The accession record and SDI are expected to be combined operations for both books and documents in which the updating information for the bibliography master for books and the bibliography master for documents will be brought together so the user has the benefit of the maximum amount of incoming information. Financial summaries are not required for documents, but this detail is replaced by further complications in the area of security control and reproduction.

SLIDE The third major area, periodical related activities, is different from the other two. Here we expect the man to stay up to date in his field by perusing journals routed to him or by examining abstract lists, such as Nuclear Science

Abstracts or International Aerospace Abstracts, etc. Consequently, you will see that the products we expect to obtain are lists of titles which can be made available to the patron, route lists which will be used by library personnel to serve the patron, and what appears to be a rather simple, but what is actually a very complicated financial and housekeeping package involved in the reordering and the recording by machine of the material obtained. In this instance, like in the book and document areas, an in-process file is necessary because as one removes material, adds titles, or takes any action which changes the available number of copies or other data, the entries must not be recorded on the master record until the date the changes become effective. The routing master is conceptually similar to the inventory master used for books and documents; however, instead of a periodic circulation list, thousands of printed route slips are produced which serve both to forward copies needed by the patron and to identify on a receipt exception basis, those magazines paid for but not obtained.

I have already mentioned that the SDI and accession information from books and documents would be combined. I should have mentioned that the circulation records will also appear in a single list. There is, of course, a small amount of circulation information about periodicals which will also appear in this record despite the fact that our policy is not to lend bound journals.

SLIDE You are undoubtedly wondering how much of ALPHA I is now in effect and how far along we are in the other phases. This slide gives an indication of the status two weeks ago when the notes for this talk were assembled.

Progress has been made since then.

The entire system study of the manual method is complete.

The generalized statements of the machine methods for the overall system are being put in finished form although a number of the parts are already farther along than that.

The periodical reordering activity has been operative since June.

The book circulation has been in effect since May.

The magazine routing program is being debugged and the first trial use will occur within the next three weeks.

The binding program is being coded at this time and will go into test during November.

The book ordering and receiving operations have been turned over to programmers for detailed coding; some trial tests will be held in November.

Periodical holdings records will be added to the periodical master and summaries obtained about the first of the year.

Mechanically-produced accession lists (books only in this case) are expected la in December.

Document circulation conversion may start about the first of the year.

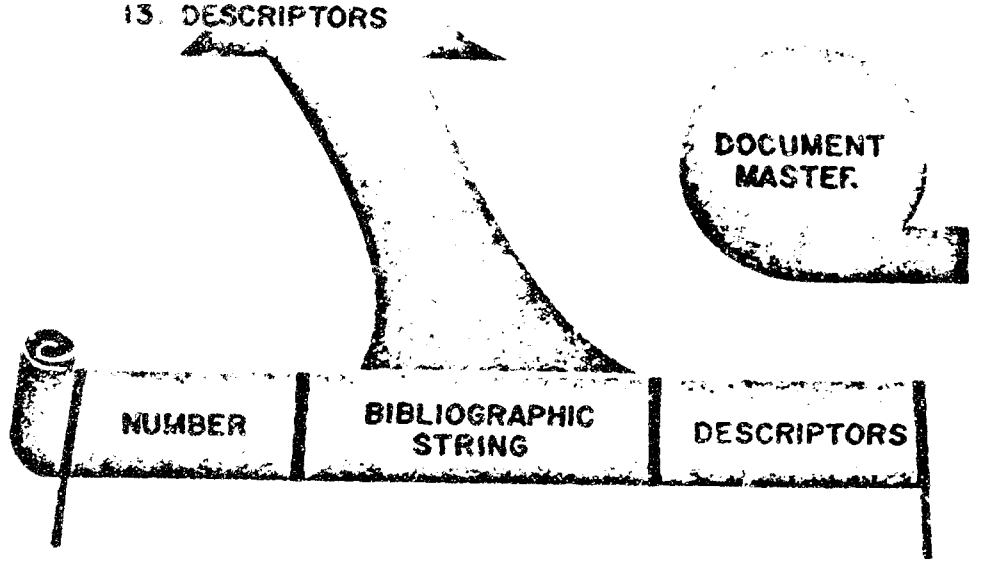
SLIDE OFF

I have had several conversations with people from other organizations who feel that the work we are doing on ALPHA I is applicable in other library situations. I agree with them and we plan to disseminate the results of our efforts in the professional press. It seems, however, that at least in the Army, there is some feeling that parts of ALPHA I may be sufficiently valuable to some other installations to use them even before they are sufficiently well tested, documented and described to put the system in the professional literature. If any of you feel you are in this position, you can write us and we will attempt to keep you informed of our progress by means of copies of progress documents we prepare. However, I must warn you that although we are

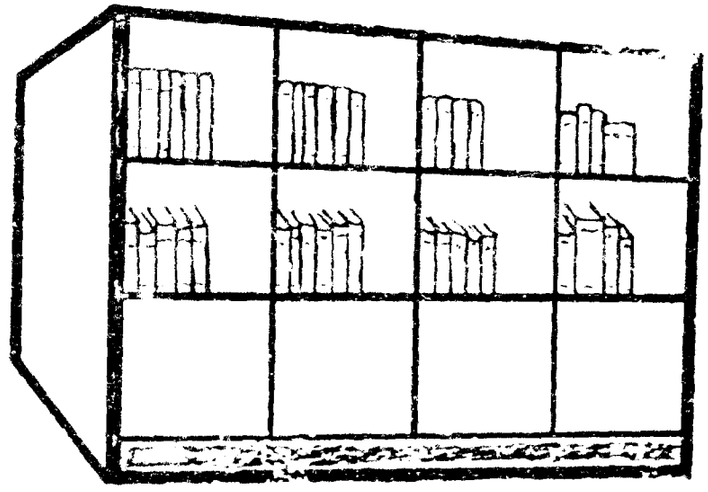
happy to share our results with you, our primary purpose at this time is to make ALPHA I the best possible system for our unique situation at Redstone Arsenal as rapidly as we can.

DOCUMENT MASTER

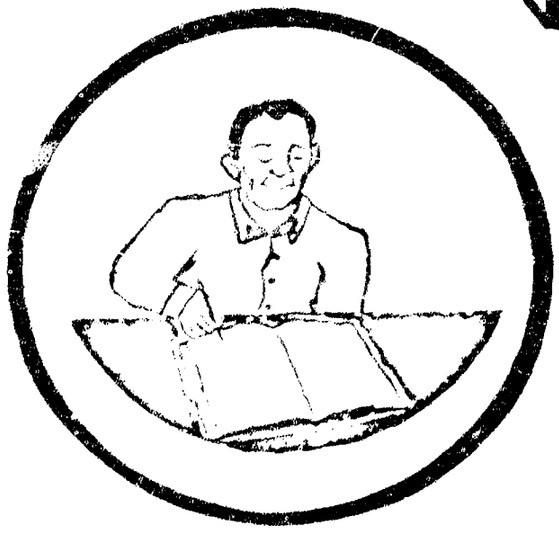
1. DOCUMENT NUMBER (includes date)
2. CORPORATE & PERSONAL AUTHOR (S)
3. TITLE (expandable)
4. ISSUING AGENCY
5. NR PAGES
6. MEDIUM (Microfilm, Roll Film, etc.)
7. ACCESSION NUMBERS
8. CONTRACT NUMBER
9. RESTRICTIONS
10. TOTAL COPIES
11. USE FREQUENCY
12. DATE OF RECEIPT
13. DESCRIPTORS



AUTOMATED
LITERATURE
PROCESSING
HANDLING AND
ANALYSIS SYSTEM
-1ST GENERATION



FILE



AUTHOR

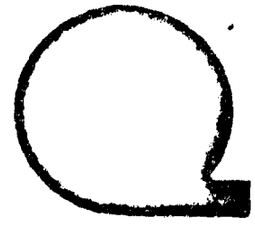
PATRON

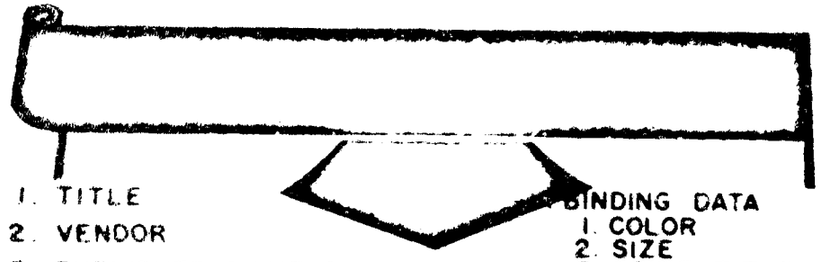
PATRON

BOOK
MASTER

DOCU[illegible]T
MA[illegible]R

PERIODICAL MASTER

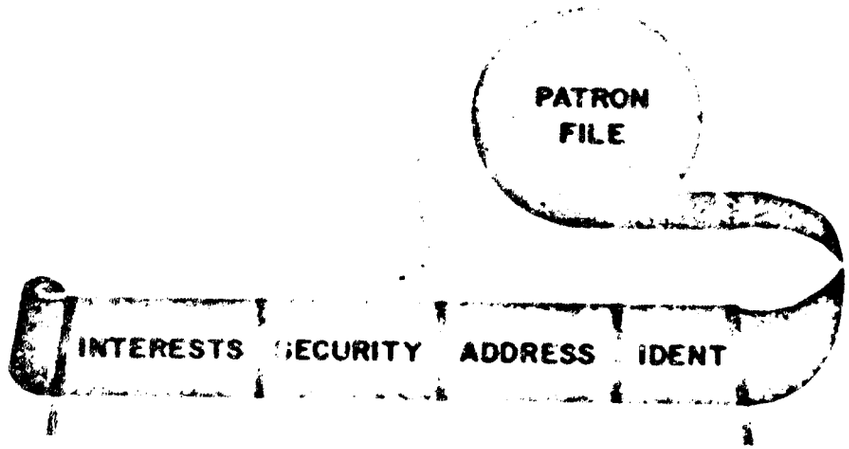




1. TITLE
2. VENDOR
3. PUBLISHER NAME & ADDRESS
4. LANGUAGE CODE
5. SUBJECT
6. FREQUENCY
7. SUBSCRIPTION DATA
   1. PRICE
   2. QUANTITY
8. BINDING DATA
   1. COLOR
   2. SIZE
   3. SPINE TITLE
   4. NOTES
   5. SCHEDULE
9. HOLDINGS
10. LACKS

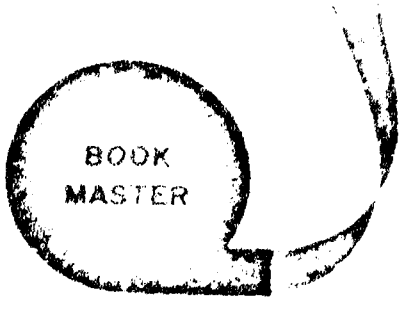
PATRON FILE

1. SOCIAL SECURITY NR.
2. NAME
3. BUILDING NR.
4. OFFICE SYMBOL
5. TELEPHONE NR.
6. PERIODICALS RECEIVED
7. "NEED TO KNOW" CODES
8. FURTHER CHARACTERIZATION
9. INTEREST DESCRIPTORS



BOOK MASTER

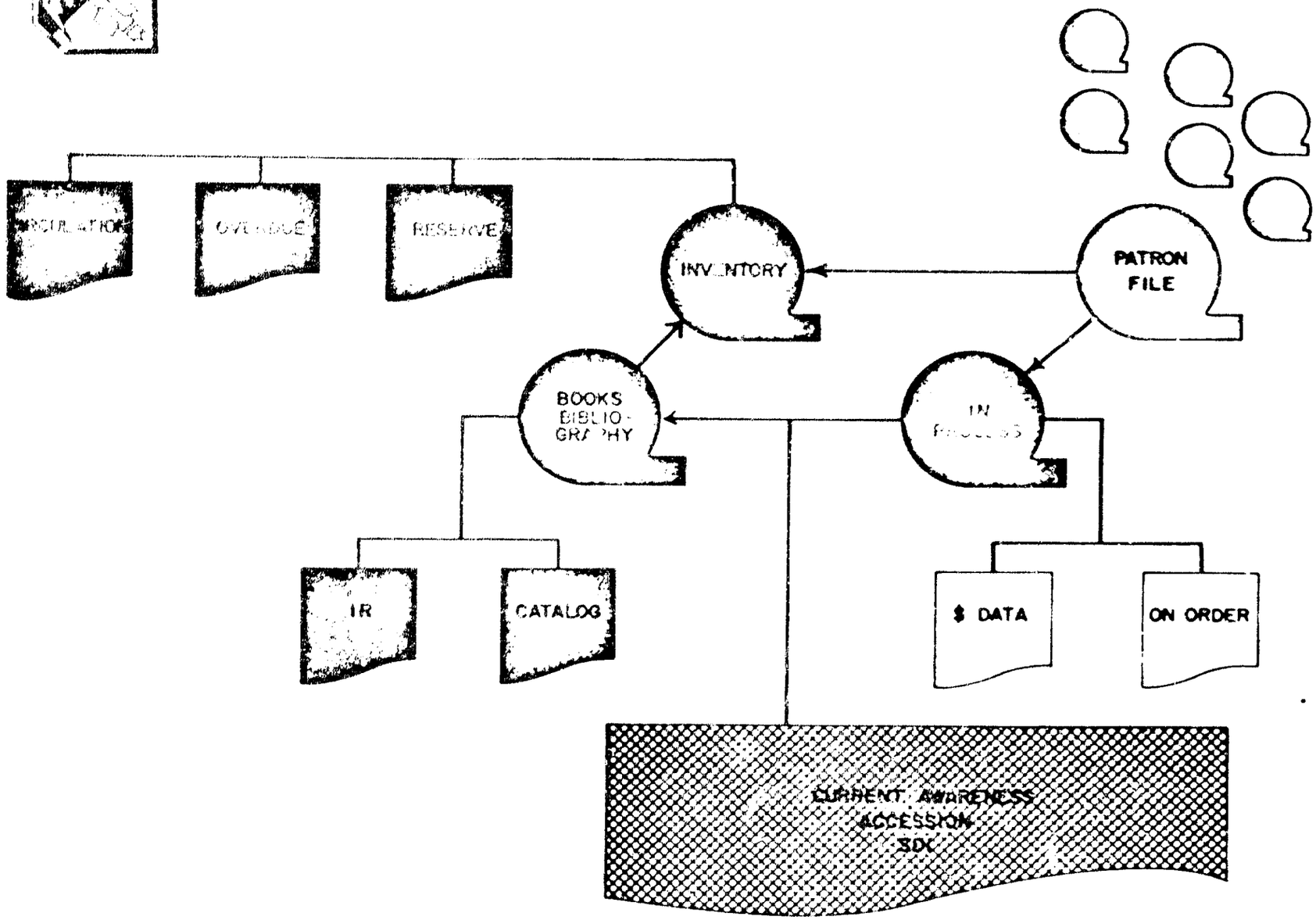
1. CALL NUMBER
2. AUTHOR (S)
3. TITLE
4. PUBLISHER & PLACE
5. DATE OF PUBLICATION
6. NR. PAGES OR VOLUMES
7. NOTES
8. NUMBER OF COPIES
9. NORMAL SUBJECT TRACINGS
10. DESCRIPTORS



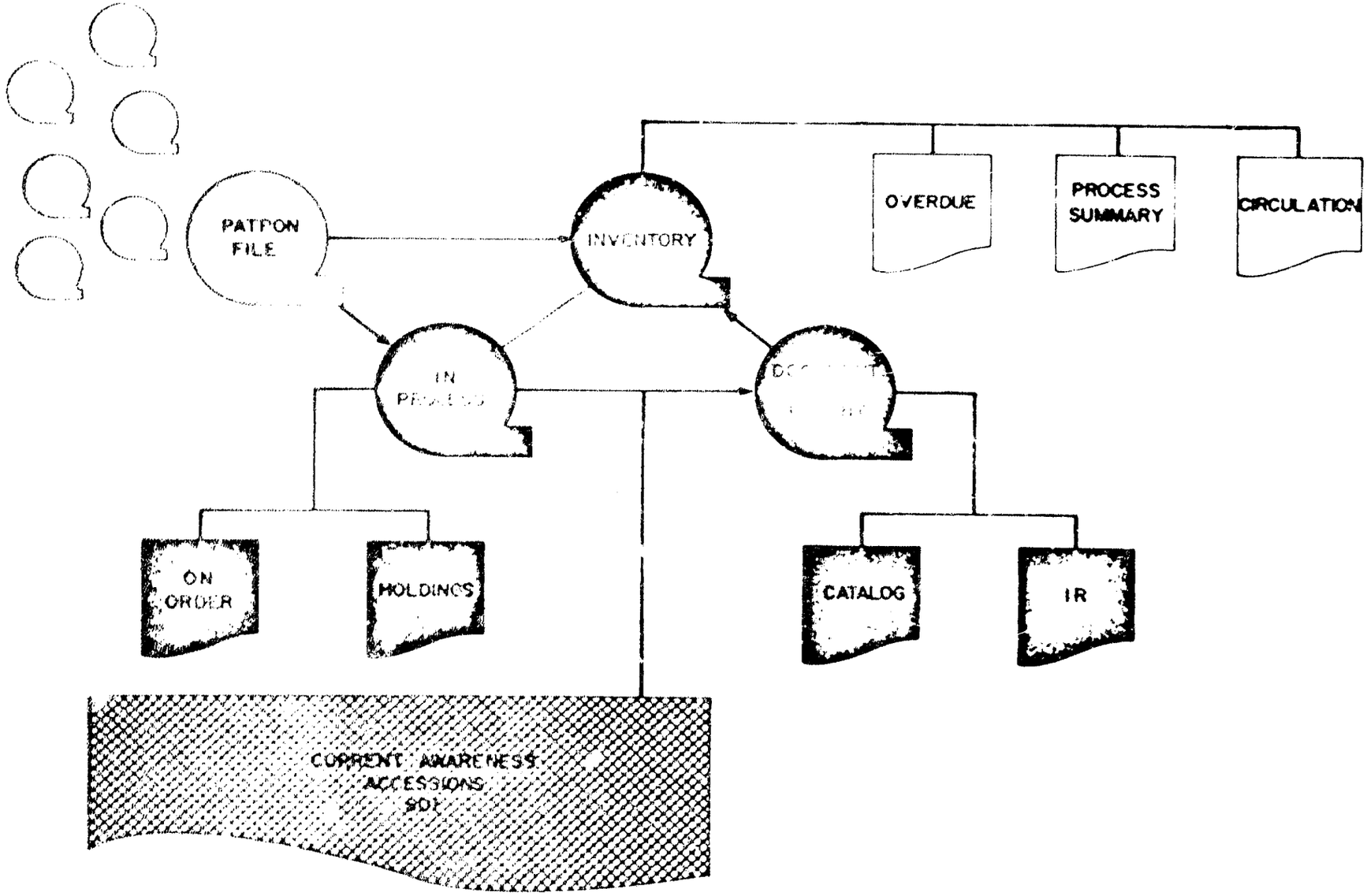
GUIDES

- EARLY MACHINE READING
- SINGLE WRITING
- GENERALIZABLE
- OPEN-ENDED
- HOPPER APPROACH
- USE WORK OF OTHERS
- IMPROVE MANUAL PROCESS

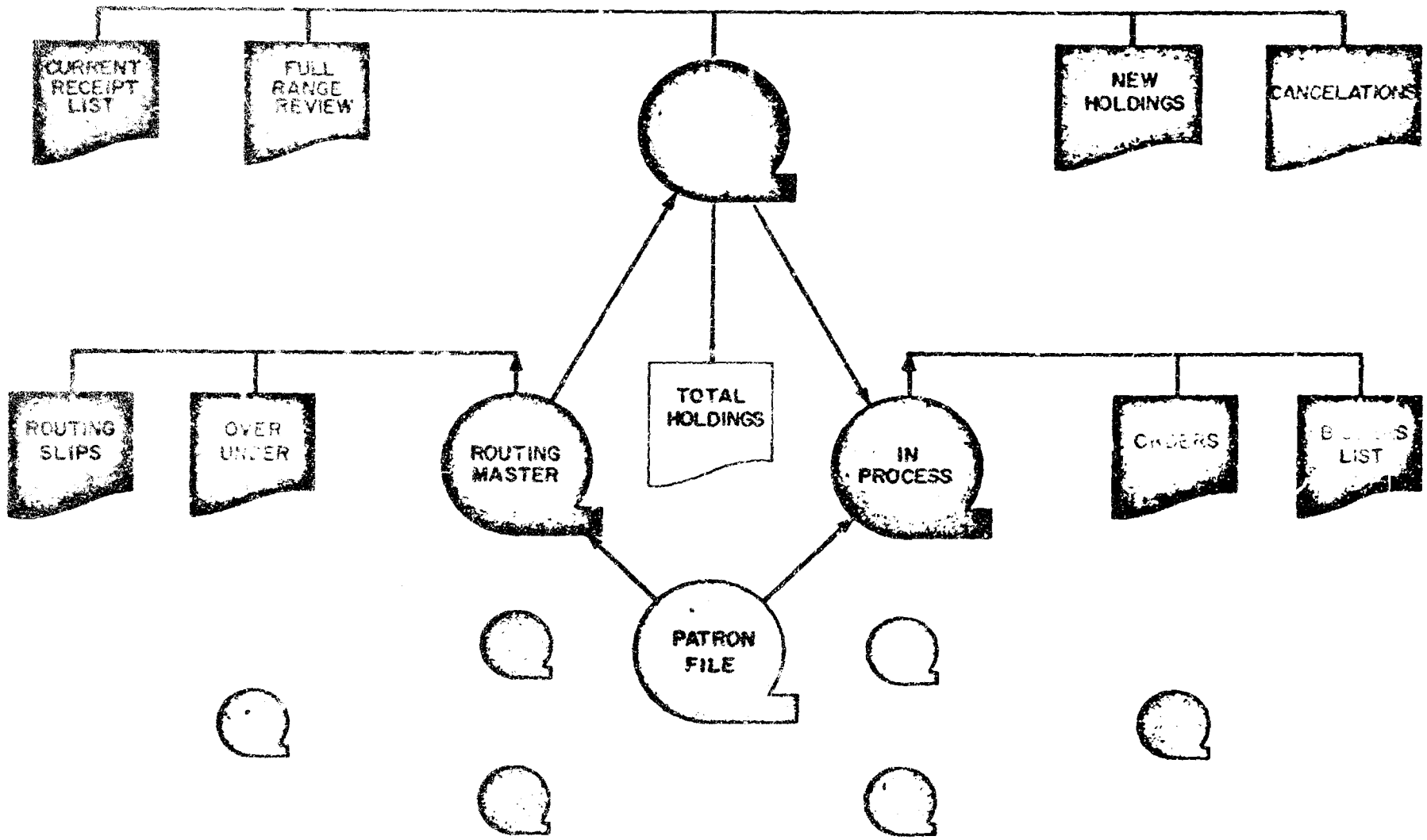
BOOKS PROCESSING



DOCUMENTS PROCESSING



PERIODICAL PROCESSING



STATUS SEP 63

OPERATING

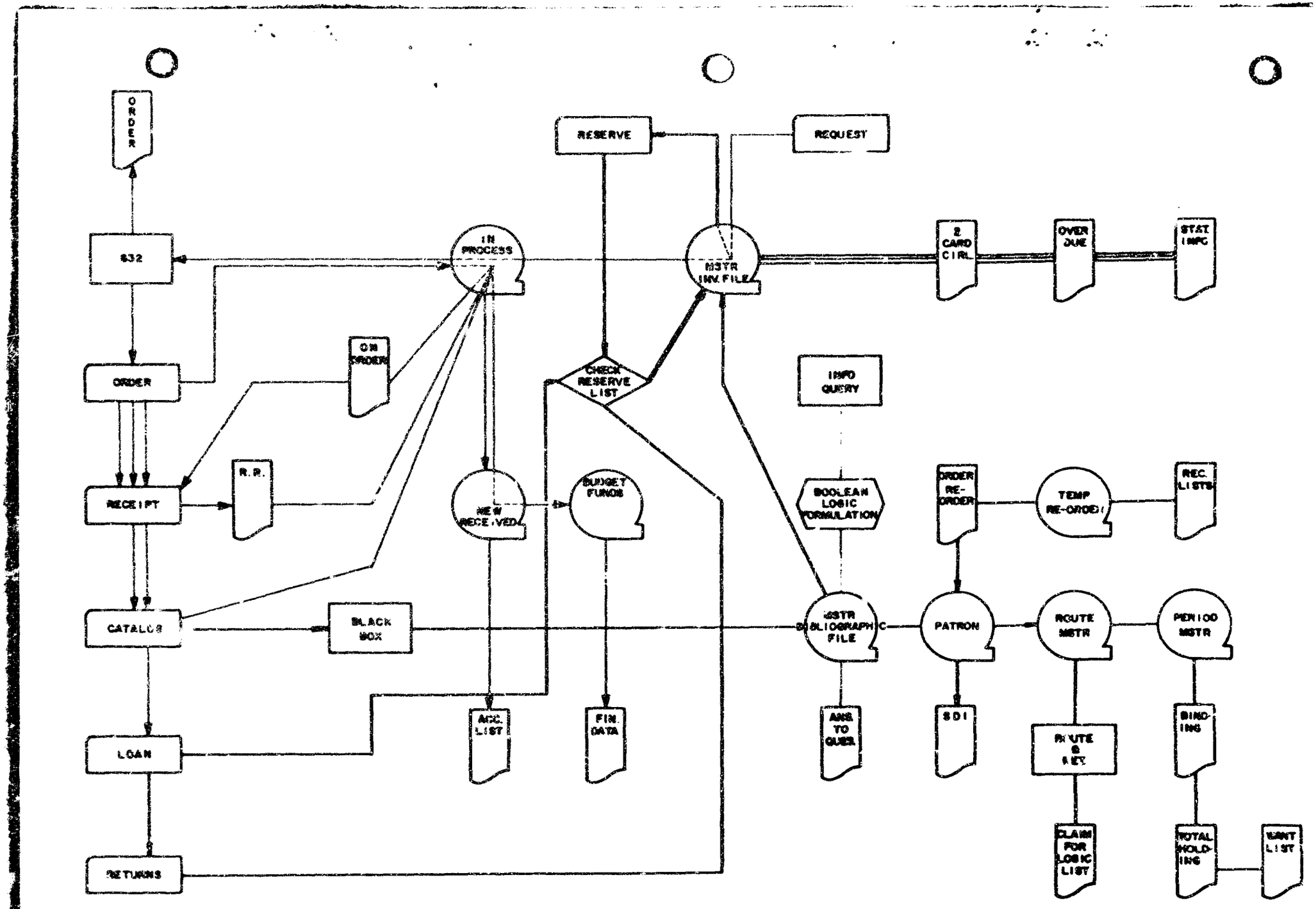
- BOOK CIRCULATION
- PERIODICAL REORDER

PROGRAM TEST

- MASTER PATRON RECORD
- PERIODICAL ROUTING
- BOOK ORDERING
- BOOK RECEIVING

PROGRAMMING

- PERIODICAL BINDING


ORDER
RESERVE
REQUEST
832
IN PROCESS
MSTR INV. FILE
2 CARD CIRC
OVER DUE
STAT IMPC
ON ORDER
ORDER
CHECK RESERVE LIST
INFO QUERY
R.R.
RECEIPT
NEW RECEIVED
BUDGET FUNDS
BOOLEAN LOGIC FORMULATION
ORDER RE-ORDER
TEMP RE-ORDER
REC. LISTS
CATALOG
BLACK BOX
MSTR BIBLIOGRAPHIC FILE
PATRON
ROUTE MSTR
PERIOD MSTR
ACC. LIST
FIN. DATA
ANS TO QUES
SDI
ROUTE & RET
BIND-ING
LOAN
CLAIM FOR LOSIC LIST
TOTAL HOLD-ING
WANT LIST
RETURNS

MILLER SEARCH ALGORITHM

1. SEARCH QUERY IS CONSTRUCTED BY FORMULAT A BOOLEAN STATEMENT USING ENGLISH TERMS AND THE LOGICAL OPERATORS *AND, OR, NOT*. THESE STATEMENTS ARE GIVEN A SEQUENTIAL PROBLEM NUMBER. STATEMENTS ARE KEYPUNCHED ONE TERM PER CARD PLUS OPERATOR, PROBLEM NUMBER, AND CARD NUMBER.

2. CARDS ARE BLOCKED ON TAPE AND SORTED BY TERMS PRODUCING THIS FORMAT.

3. LIKE TERMS ARE SUMMARIZED APPEND TO EACH TERM QUESTION NUMBER, CARD NUMBER, & OPERATOR.

4. MATCH QUESTION TERMS TO SUMMARIZED MASTER TERM FILE & APPEND REFERENCES CREATE A RECORD FOR EACH QUESTION CONTAINING ALL REFERENCES TO OCCURENCES OF THAT TERM. SORT BY QUESTION NUMBER & CARD NUMBER TO COLLECT ALL MEMBER TERMS OF A GIVEN QUESTIONS. EVALUATE BOOLEAN OPERATORS AND PRODUCE ALL "TRUES"